OECD GUIDELINES DOCUMENTS

If you require further information please contact the OECD Secretariat

PROJECT TITLE

OECD WNT Project: Construction of a series of guidance documents for consistent reporting of 'omics data from various sources

SUBMITTED BY (Country / European Commission / Secretariat)

Main Lead: UK/PHE/Health Canada

DATE OF SUBMISSION TO THE SECRETARIAT

June 2017

DETAILS OF LEAD COUNTRY/CONSORTIUM

Country /Organisation:	UK/PHE
Agency/ministry/Other:	Public Health England/Department of Health
Mail Address:	Centre for Radiation, Chemical and Environmental Hazards, PHE, Chilton, Oxon OX11 0RQ. UK
Phone/fax:	01235 825139
Email:	Tim.Gant@phe.gov.uk Carole.yauk@canada.ca thomas.russell@epa.gov m.viant@bham.ac.uk cc. Miriam.Jacobs@phe.gov.uk

PROJECT OUTCOMES

☐	New Test Guideline	☒	Guidance document
☐	Revised Test Guideline	☐	Detailed Review Paper
☐	Deletion of an existing Test Guideline		Other, please specify below

---

Background

High density data generating 'omics technologies have been an integral part of toxicology for about 20 years. Toxicology research using these methods has contributed to the greater understanding of modes and mechanisms of toxicity, and the identification and development of biomarkers of toxicity. Additional proposed applications in regulatory toxicology include read-across and identification of point of departure. However, toxicogenomics has generally failed to live up to expectations in regulatory application as an endpoint assay for adverse effect (e.g., application of a gene expression profile indicative of an impending pathological event) or hazard profiling. In failing to achieve this, some early expectations, such as reducing and refining animal testing, have not been realised to the degree initially thought possible.

Several reasons for this lack of regulatory penetrance in applicability are likely including: 1) poor experimental design and data quality in early studies led to lack of reproducibility and a bad reputation for various omics technologies; 2) lack of accepted quality control standards and data quality assessment tools during data generation, 3) lack of availability of the required metadata associated with the 'omics data that are necessary for interpretation and regulatory application, 4) a lack of transparency, public availability and best practices/standards for the data processing methods used to turn raw data into an interpretable list of observations, 5) variances in methods and data used to analyse and interpret genomics data, and 6) lack of standardized reporting frameworks or templates to ensure that all of the required and appropriate data, associated metadata, and analytical processes are readily available.

Across the various 'omics fields, the area of transcriptomics has progressed furthest in development. For transcriptomics there have been more international activities that have, and are, addressing some of the obstacles and application of the data. For example, quality of data generation both for microarrays and RNA sequencing was the subject of a series of papers from the Microarray Quality Control Consortium [1-11], and best practices and standards have been proposed from this work. 'Omics associated metadata was the subject of an early paper (2001) from Brazma *et al.* [12] that led to the adoption of standards for the submission of the data to the two major public repositories (the Gene Expression Omnibus and ArrayExpress). The requirement of journals for data to be deposited in these repositories before publication has ensured that at least some relevant data/metadata are available. For metabolomics, there was a considerable focus on establishing reporting standards by the international community – termed the Metabolomics Standards Initiative (MSI) – leading to multiple publications in 2007 [13]. As with transcriptomics, there has been an increasing requirement by journals and international data repositories (e.g. MetaboLights and Metabolomics Workbench) to comply with metabolomics data and metadata standards.

The European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) held a workshop last year (Workshop on applying 'omics technologies in chemicals risk assessment, Madrid, Spain, October, 2016) to directly discuss how to overcome these obstacles that limit regulatory uptake of 'omics data. The workshop discussions confirmed that best practices when performing 'omics studies for regulatory purposes will be important for increasing regulatory uptake. However, participants felt that prescriptive guidances/protocols for specific 'omics data processing technologies might not be helpful. Tools and analytical pipelines in these areas are constantly evolving/improving, and fit-for-purpose approaches are usually required and experiment-dependent. Instead, as a contribution to establishing best practices for 'omics studies in a regulatory context, reporting frameworks were suggested as a way forward, which included: (i) Good-Laboratory Practice-like context for collecting, storing and curating 'omics data; (ii) the processing of 'omics data, and (iii) weight-of-evidence approaches for integrating 'omics data. Meeting participants believe that these frameworks (Figure 1) will contribute to the establishment of a baseline for best practice of 'omics studies, thereby facilitating the regulatory applicability and use of 'omics studies. Building on this momentum, an ECETOC supported project led by the UK (Viant, Univ. Birmingham and Ebbels, Imperial College London) has been started specifically with the objective to develop best practice and reporting standards for metabolomics applications in regulatory toxicology (MEtabolomics standaRds Initiative in Toxicology; MERIT).

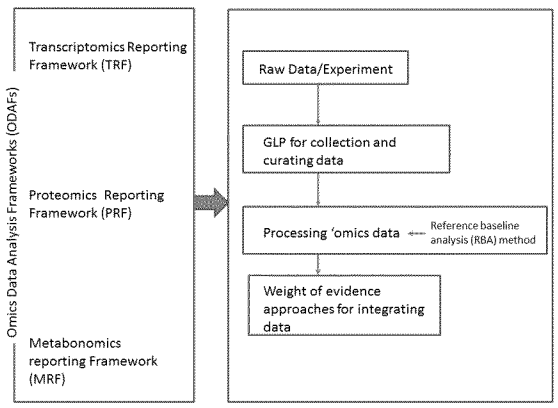

Figure 1 - Reporting Frameworks

In this OECD project we propose to focus initially on the area of transcriptomics (starting in the summer of 2017) leading to the development of a Transcriptomics Reporting Framework (TRF). This TRF leverages the significant advances made by ECETOC in developing a discussion/concept document for consideration by OECD EAGMST partners on best practices in reporting transcriptomics data and development of either a guidance or harmonized reporting and analysis template for transcriptomic data (e.g., microarray, QPCR and RNA-seq). Following the progress made in the ECETOC MERIT project, and building on the early work of the TRF template, this OECD project will expand to include metabolomics (MRF: estimated to start in summer 2018).These TRF and MRF activities can then be used as templates for developing reporting frameworks in other 'omics areas such as proteomics.

If 'omics methods are to fulfil their promise of two decades ago in improving testing and reducing animal use in toxicity testing, then reporting frameworks and best practices (including relevant quality standards) must be developed to facilitate their use. This is an important first step towards achieving those objectives.

Bibliography

1. Wang, C., et al., *The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance.* Nat Biotechnol, 2014. **32**(9): p. 926-32.

2. Liu, Z., et al., *Comparative Analysis of Predictive Models for Nongenotoxic Hepatocarcinogenicity Using Both Toxicogenomics and Quantitative Structure–Activity Relationships.* Chemical Research in Toxicology, 2011. **24**(7): p. 1062-1070.
3. Chen, M., et al., *Selecting a single model or combining multiple models for microarray-based classifier development?--a comparative analysis based on large and diverse datasets generated from the MAQC-II project.* BMC Bioinformatics, 2011. **12 Suppl 10**: p. S3.
4. Shi, L., et al., *The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models.* Nat Biotechnol, 2010. **28**(8): p. 827-38.
5. Luo, J., et al., *A comparison of batch effect removal methods for enhancement of prediction performance using MAQC-II microarray gene expression data.* Pharmacogenomics J, 2010. **10**(4): p. 278-91.
6. Fang, H., et al., *Arraytrack: an FDA and public genomic tool.* Methods in Molecular Biology, 2009. **563**: p. 379-398.
7. Shi, L., et al., *Reproducible and reliable microarray results through quality control: good laboratory proficiency and appropriate data analysis practices are essential.* Curr Opin Biotechnol, 2008. **19**(1): p. 10-8.
8. Fuscoe, J.C., W. Tong, and L. Shi, *QA/QC issues to aid regulatory acceptance of microarray gene expression data.* Environ Mol Mutagen, 2007. **48**(5): p. 349-53.
9. Sirotnak, F.M., et al., *Co-administration of probenecid, an inhibitor of a cMOAT/MRP-like plasma membrane ATPase, greatly enhanced the efficacy of a new 10-deazaaminopterin against human solid tumors in vivo.* Clinical Cancer Research, 2000. **6**(9): p. 3705-3712.
10. Berman, E., M. McBride, and W. Tong, *Comparative activity of tamoxifen and N-desmethyltamoxifen in human multidrug resistant leukemia cell lines.* Leukemia., 1994. **8**(7): p. 1191-1196.
11. Fan, X., et al., *Consistency of predictive signature genes and classifiers generated using different microarray platforms.* Pharmacogenomics Journal. **10**(4): p. 247-257.
12. Brazma, A., et al., *Minimum information about a microarray experiment (MIAME) -toward standards for microarray data.* Nature Genetics, 2001. **29**(4): p. 365-371.
13. Multiple papers dedicated to reporting standards, *Metabolomics* journal, 2007, volume 3.

Progress

A proposal was brought to the EAGMST meeting of June 2016 to develop a guidance document to deal with just the processing of data from collection to interpretable list. The meeting, however, favoured expanding this in to a series of guidance documents dealing with all steps in an omics pipeline, and requested a sub-group be formed to develop a scoping document for the work areas to be presented at the December EAGMST meeting.

The subgroup met in Madrid in October 2016 as part of the ECETOC meeting on 'Applying Technologies to Risk Assessment' and consisted of:

Tim Gant (UK), Matthew Martin (USA), Aldert Piersma (Netherlands), Carole Yauk (Canada), George Daston (USA), David Rouquie (France), Herve Seitz (France), Weida Tong (USA), Ben van Ravenzwaay (Germany).

Milestones

The group agreed a path forward:

1) Under the auspices of ECETOC develop a paper that sets out the background and need for Guidance Document/reporting frameworks for the processing of transcriptomics data for regulatory use.
2) Develop a draft guidance document that explores the data collection and processing work area in toxicogenomics from the output of the ECETOC meetings of June 2015 (Brussels) and October 2016 (Madrid). This guidance document will contain proposals for: (i) Good-Laboratory Practice-like context for collecting, storing and curating 'omics data; (ii) the processing of 'omics data (with a focus on transcriptomics) including a Reference Baseline Analysis (RBA) method; and (iii) weight-of-evidence approaches for integrating 'omics data.
3) Publish 1 and 2 as an ECETOC report in a special journal issue.

Time	**Activity**	**Status**
July 2015	First consortium meeting to define the framework and terms of reference	Complete
Sept 2015	First presentation of the framework at EUROTOX	Complete
Nov 2015	Presentation of the first consortium meeting output at the LRI meeting in the form of a poster	Complete
January 2016	Second consortium meeting to produce a final draft of the framework	Complete
April 2016	Presentation of the draft framework to ECHA	Complete
June 2016	First presentation at EAGMST meeting	Complete
July 2016	Third consortium meeting to consider the outcomes of the EAGMST meeting	Complete
October 2016	Discussion and consideration by EAGMST sub-group at ECETOC meeting 'Applying Technologies to Risk Assessment' – development of the ORF concept.	Complete
December 2016	Presentation of the scoping document for EAGMST	Complete
March 2017	Presentation at SOT in a dedicated workshop: Data Standardization Across 'Omic Platforms in Regulatory Toxicology	Complete
Aug – June 2017	ECETOC led development of a background paper that explores the challenges for using 'omics data in regulatory submission and need for guidance documents for 'omics data analysis and interpretation	Underway
April 2017	Launch of ECETOC MERIT project focused on best practice and reporting standards in regulatory metabolomics	Underway
June 2017	Presentation of the outcome of the ECETOC workship and the concept of 'omics reporting frameworks to EAGMST: Decision to focus on two reporting frameworks (TRF and MRF) and a 'Reference Baseline Analysis' project	Complete
July 2017	Coordination of core leadership team	Complete
Sept 2017	Presentation at the 53rd EUROTOX Congress	Sept 2017
Sept. 2017	Creating of TRF working group and project initiation	
June 2018	Expansion of MERIT project to OECD Metabolomics Reporting Framework (MRF) project	June 2018
End-	Publication of TRF and MRF	

deliverable		

2. Will additional information, including generation or collection of data, be required? If yes, please describe the anticipated process and timelines.

ECETOC are supporting the development of frameworks designed to facilitate the transfer of 'omics methodology from the research to the regulatory sphere including: *(i) for establishing a GLP-like context for collecting, storing and curating 'omics data; (ii) for the reporting of 'omics data; and (iii) for the application of WoE approaches in the interpretation of 'omics data*. In this project we are just considering ii) the reporting of 'omics data. This includes within it the work presented at EAGMST in June and December 2016 that had a narrower focus of the development of a Reference Baseline Analysis (RBA). This RBA is now included in the reporting frameworks as a means of developing a reference baseline set of differentially expressed genes (DEGs).

3. Indicate the estimated overall resource need (time/money) for member country / consortium and Secretariat

There are no sources in the current document.

This is an international consortium (see names under Progress above), financially sponsored by CEFIC under the auspices of ECETOC.

Financial support includes travel and resources for data generation and analysis.

Industry

CEFIC (Belgium) will be investing financial resource for travel support of the consortium and the generation of data.

ECETOC (Belgium) will be investing personnel resources to manage the consortium

BASF (Germany) has produced data for to inform the analysis under the framework by the consortium.

BAYER (France) will be investing time in the development of the framework, writing and analysis of data and attendance at meetings

Nestlé (Switzerland) will be inversing personnel time in the development of the framework in attending meetings, contributions and drafting of the framework.

SAS analytics (USA) will be investing time in the analysis of data and contributing to the development of the framework.

UK

The UK (PHE/QUB/Imperial/Birmingham) will be investing time in the development of the TRF and MRF frameworks, writing and analysis of data and leading the consortium.

USA

The US FDA will be investing time in the development of the framework, writing and analysis of data and attendance at meetings

Canada

Canada will be investing time in the development of the framework, writing and analysis of data, and meeting attendance.

China

Investing academic resources from the University of Shanghai.

France

Nice University will be investing personnel resource in the development of the framework and attendance at meetings.

4. Is this proposal intended to replace an existing Test Guideline or lead to the deletion of an existing Test Guideline?

No.

This is a new guidance document designed to provide a better foundation for the reporting of 'omics data for use in regulatory submissions.

ESSENTIAL INFORMATION

In this section, please provide the information required by the Working Group of National Coordinators of the Test Guidelines Programme to assess the suitability of the project for the workplan of the Test Guidelines Programme

1. What is the existing or expected regulatory need/data requirement that will be met by the proposed outcome of the project? Please provide details below or as an attachment.

This project aims to develop reporting frameworks for the standardisation of reporting of 'omics data generation and analysis, to ensure that all of the information required to understand, interpret and reproduce an 'omics experiment and its results are reported. The TRF includes a RBA that provide a prescriptive analysis method to recognise DEGs and acts as a baseline comparator dataset against which other operator methods can be compared and allows cross comparison between studies. Additionally the reporting frameworks include an equivalent activity for metabolomics, forming the MRF.

The reporting framework will require the inclusion of information associated with experimental variables and methodologies, including specific sample handling information, wet laboratory methodological details, and data extraction, manipulation and analysis pipelines for 'omics experiments. It is not the purpose of the reporting framework to stipulate the methods of data analysis or interpretation, only to ensure that sufficient information is available to enable an evaluation of the quality of the experimental data and interpretation, and analysis reproducibility. The TRF though will contain a RBA to be used to develop a reference baseline set of data against which comparisons can be made.

This work will build on other work that defined reporting criteria for the deposition of 'omics data in public repositories, and on quality control assessment and reproducibility of 'omics data.

This framework is required to advance the use of 'omics data in regulatory toxicology, through increasing transparency in how 'omics data were derived and analysed, facilitating increased reproducibility of the 'omics analyses. Though designed specifically to assist the regulatory community it is anticipated it will also find application in research.

or as attachment No.__

2. How will the work contribute to further international harmonisation of hazard and risk assessment? Please provide details below or as an attachment.

It is anticipated that application of the framework will enhance the ability to use 'omics data in regulatory assessments through increasing: (1) understanding by the evaluator of how the 'omics data were generated, (2) ability to assess data quality, and (3) reproducibility of data analysis in the 'omics area by fully reporting on the 'omics analytical pipeline applied. In addition, the RBA will be used to derive a reference baseline set of data against which comparisons can be made. This will provide a better foundation for the use of these data in the assessment of hazard and risk.

3. How will the proposed project address issues and /or endpoints that are of major human health or environmental concerns? If there are existing Test Guidelines or projects in the work plan of the Test Guideline Programme covering the same endpoint, please refer to these and describe the added value and usability of the proposed new test method. Please provide details below or as an attachment.

To our knowledge there are no such guidance documents published by the OECD in recent years, and no such projects on the OECD rolling workplan. 'Omics data provide a powerful means to explore the broad biological impacts of a toxicological exposure, through assessment of thousands of endpoints in parallel. The molecular perturbations can be measured in vitro (and thus the work aligns with projects in the area of NAMs and IATA) or *in vivo*, and in any experimental model (and thus is relevant to both human and ecological health). 'Omics methods greatly facilitate understanding of mode of action and human relevance. Therefore, there is significant motivation to move the use of 'omics from research into regulatory areas. Acceptance of 'omics data though is hindered by the complexity of the data that prevents transparency in analysis and therefore the derivation of conclusions.

or as attachment No.__

4. Will the project have general support from OECD member countries or is the outcome relevant for just one or a few member countries / stakeholders? Provide details of the countries and the rationale for this view below.

☒ Many countries ☐ A few countries ☐ Only for the submitting country

Many countries are involved in the drafting (see above) and it will have relevance for all.

5. If the Test Guideline is not intended for general use, indicate if the Test Guideline would be intended for:

☐Specific (limited) applications such as pesticide usage, or

☒for specific classes of chemicals (e.g. surfactants) rather than for chemicals in general.

6. If the expected outcome of this proposal is a Test Guideline or a Guidance Document, provide information on the intended use, applicability and limitations of the test method.

The intended output is a guidance document and not a test guideline.

7. Provide supporting information on the validation status (i.e. relevance and reliability) of the method. Principles for validation of test methods for OECD Test Guidelines are described in Guidance Document 34.

Provide justification and rationale for the test, including data.

If there are no or limited data available to support the reliability and relevance of the proposed test, indicate if validation work is included in the project.

If there is no need for validation, provide a detailed justification.

This is no test method or test guideline and therefore validation is not necessary *per se*. However, limited validation of the reporting frameworks will be undertaken in the process of its generation using pre-existing data suitable for the purpose, and through case studies on use and application.

8. Describe if the test method includes components, equipment or other scientific

procedures that are covered (or pending) by Intellectual Property Rights (IPR) (e.g., patents, patent applications, industrial designs and trademarks). Information should be provided on the overall availability of the IPR-protected components including whether they are commercially available or require a Material Transfer Agreement (MTA) or other licensing agreements. In addition, the possibility of providing a generic description of the IPR-covered component/test system should be disclosed and whether Performance Standards have been developed for the test method.

No parts of the TRF, MRF or RBA will be covered by IPR.

ADDITIONAL INFORMATION

In this section please provide further information to allow the Working Group of National Coordinators of the Test Guidelines Programme to assess the suitability of the project for the workplan of the Test Guidelines Programme

1. If the expected outcome of the project proposal is a Test Guideline and is based on existing, regional or international documents such as guidelines, protocols or guidance material, please provide that information here or as an attachment.

The guidance document will build on existing work that has been carried out by the Microarray Quality Control Consortium (MAQC) and the international Metabolomics Standards Initiative (MSI) that has been published in the open literature. The MAQC and MSI have produced guidelines for data generation to ensure the quality of the raw data. Part of these reporting frameworks will specify that these guidelines for data generation should be adhered to. The frameworks will then specify the experimental details that must be provided during a regulatory submission of 'omics data and provide a RBA method for the recognition of DEGs that are then taken for interpretation. The RBA method is prescriptive, but designed not to be an end point in its own right. Instead, the RBA provides a baseline reference data set against which comparisons can be made.

or as attachment No.__

2. If Animal Welfare considerations are addressed in the project proposal, provide details below or as an attachment. Explain if the project is aimed at refining, reducing and/or replacing the use of animals.

If the project is not specifically developed for animal welfare purposes, indicate if the animal welfare considerations have been a component of the project proposal.

Indicate if animal welfare considerations are irrelevant to the project, for example for physico-chemical properties.

Animal welfare issues are not relevant to the document. This work will have been carried out in accordance with all relevant national law and with respect for the welfare of the animals.

or as attachment No.__

3. Provide information on expected or possible resource savings in member countries as a result of this project.

A standardised TRF and MRF will increase the utility of these 'omics data in the hazard assessment of chemicals, in assessing human relevance, and for a variety of applications.

4. If the expected outcome of the proposed project is a Guidance Document or Detailed Review Paper, will it be directly linked to the development of a particular Test Guideline or a series of Test Guidelines?

☐ Yes, it is the initial step in the development of a new or revision of existing Guidelines.

☐ Yes, additional guidance is needed for the most appropriate selection of the Guidelines on the subject.

☒ No, the guidance is on issues related to testing or the development of Test Guidelines in general.

There are _0_ attachments added to this form.

ASSESSMENT OF PROJECT PROPOSAL

(To be completed by all member countries /stakeholders except the submitter)

Country / Organisation:	

Representative: (Preferably NC):	

Taking into account the project information, requested above, does this project meet the needs of the member countries for addition to the workplan of the Test Guidelines Programme

☐ Yes ☐ No ☐ Further information needed

If the response is "No" or "Further information needed", please provide justification:

Remarks as appropriate, including further information needs, if any: